arXiv:1211.2438v1 [math.DS] 11 Nov 2012

EXPLICIT CORRELATION BOUNDS FOR EXPANDING CIRCLE MAPS USING THE COUPLING METHOD

HENRI SULKU

Abstract. In this paper, several fundamental facts, especially the existence and uniqueness of an absolutely continuous ergodic measure with an exponential mixing rate, are derived for smooth expanding circle maps. Although the results are classical, the coupling method is relatively modern and employed here for the first time to yield explicit bounds in terms of system constants. The work constitutes a part of the author's Bachelor's thesis. The author hopes that the manuscript can serve as a useful introduction to the flexible coupling method in the theory of dynamical systems.

Acknowledgements. The thesis was written under a supervision of Dr Mikko Stenlund and I use this opportunity to express my gratitude for his support and professional guidance during the process.

1. Introduction

We assume that $T : \mathbb{S}^1 \to \mathbb{S}^1$ is a smooth uniformly expanding map on the circle. More precisely, let $\mathbb{S}^1$ denote the interval $[0, 1]$ with its endpoints identified, let T be C^2, and assume there exists $\lambda > 1$ such that

$$|T'(x)| \geq \lambda \qquad \forall\, x \in \mathbb{S}^1. \tag{1}$$

Our goal is to prove that there exists an invariant measure[1] which is

(i) absolutely continuous with respect to the Lebesgue measure and

(ii) mixing (thus also ergodic).

Recall that a measure is invariant if $\int_{\mathbb{S}^1} f \circ T \,\mathrm{d}\mu = \int_{\mathbb{S}^1} f \,\mathrm{d}\mu$ for all integrable functions f. It is called mixing, provided that

$$\int_{\mathbb{S}^1} f \circ T^n g \,\mathrm{d}\mu \longrightarrow \int_{\mathbb{S}^1} f \,\mathrm{d}\mu \int_{\mathbb{S}^1} g \,\mathrm{d}\mu \qquad \forall\ f, g \in L^2(\mu).$$

Note that here $\int_{\mathbb{S}^1} f \circ T^n \,\mathrm{d}\mu = \int_{\mathbb{S}^1} f \,\mathrm{d}\mu$ due to invariance. Thus, the mixing condition above has the probabilistic interpretation that the random variables $f \circ T^n(x)$ and $g(x)$ become asymptotically de-correlated as $n \to \infty$, if x is distributed according to μ.

Actually, we want to prove stronger statements than these and obtain quantitative information about the invariant measure. In addition to (i) and (ii), we are going to show that the invariant density (the density of the invariant measure) is actually Lipschitz continuous and that it is exponentially mixing in a suitable class of observables; namely, the class of Hölder continuous functions of arbitrary order. Our main result is thus the following:

2000 *Mathematics Subject Classification.* 37D20.

Key words and phrases. Expanding maps, coupling, correlation decay.

[1] A natural choice for the governing σ-algebra is the Borel σ-algebra $\mathcal{B}$.

Theorem 1.1. *An invariant measure μ satisfying properties (i) and (ii) above exists. In fact, μ is the only ergodic measure satisfying (i). Additionally, for bounded f and Hölder continuous g,*

$$\left| \int_{\mathbb{S}^1} f \circ T^n g \,\mathrm{d}\mu - \int_{\mathbb{S}^1} f \,\mathrm{d}\mu \int_{\mathbb{S}^1} g \,\mathrm{d}\mu \right| \leq C \|f\|_\infty (\|g\|_\infty + H_\alpha(g)) \theta^{\alpha n}, \tag{2}$$

with

$$C = 96(2 + \Omega) \exp(\exp(4(\Omega + 1))) \tag{3}$$

and

$$\theta = (1 - \exp(-3(\Omega + 1))^{(\log(\lambda)/(4(\Omega+1)))} \in (0, 1). \tag{4}$$

Here $H_\alpha(g)$ stands for the Hölder constant of a function g.

Structure of the paper. The treatise, and the proof of 1.1 as well, is divided into five sections: first we express some preliminary results which will play a central role in the study of the system. Then, in the third section, we prove the existence of an invariant measure with a Lipschitz continuous density. The coupling argument is employed in the fourth section where we prove the exponential decay of correlation coefficients for Hölder continuous observables. Some of the properties are studied in a bit more general setting in [1] as well. Finally, at the fifth section we prove the mixing property by extending the correlation decay from C^α to $L^1(\mu)$. Remark that the exponential rate is not preserved but even the continuous functions can be too irregular instead. Additionally, a couple of fundamental results related to the topic of the work are presented in the appendix because those are somewhat isolated from the general storyline.

2. Preliminaries

We begin by introducing some basic facts that will be used throughout the text. The expansion property below is a manifestation of the hyperbolic nature of the dynamics and inverse branches are needed to treat the transfer operator effectively. The transfer operator itself will be a crucial tool in the analysis of the dynamics.

2.1. **Expansion properties and inverse branches.** First, since T is C^2 and (1) holds, T' cannot change sign. Therefore, we can and do assume that T' is strictly positive:

$$T'(x) \geq \lambda \qquad \forall\, x \in \mathbb{S}^1. \tag{5}$$

We denote the standard metric on $\mathbb{S}^1$ by $d(\,\cdot\,,\,\cdot\,)$, i.e., $d(x,y) = \min\{|x-y|, 1-|x-y|\}$ and arc lengths by $|\cdot|$. Observe that, given an arc $I \subsetneq \mathbb{S}^1$ such that $T(I) \neq \mathbb{S}^1$, it is obviously true by the expanding property (5) that $T(I)$ is an arc with $|T(I)| \geq \lambda |I|$. Notice that the quantity $w = \int_{[0,1]} T'(x)\,\mathrm{d}x$ is a positive integer. We present the following, rather obvious, result without proof.

Lemma 2.1. *Let $J \subsetneq \mathbb{S}^1$ be an arc. Then the map T has exactly w well-defined inverse branches on J. More precisely, there exist arcs $I_i \subsetneq \mathbb{S}^1$ and maps $T_i^{-1} : J \to I_i$, $1 \leq i \leq w$, such that each restriction $T|I_i$ is a one-to-one and onto J (in fact a C^2 diffeomorphism in the interior of I_i) and T_i^{-1} is its inverse.*

By the earlier observation, $|J| = |T(I_i)| \geq \lambda |I_i|$. Therefore, Lemma 2.1 can be applied repeatedly, resulting in exactly w^n well-defined inverse branches of T^n on J. Let us denote them $(T^n)_i^{-1} : J \to I_i^n$, $1 \leq i \leq w^n$, where each I_i^n is an arc. Iterating the previous bound, $|I_i^n| \leq \lambda^{-n}|J|$. This results in the lemma below.

Lemma 2.2. *Let $x, y \in \mathbb{S}^1$ be arbitrary. Suppose $J \subset \mathbb{S}^1$ is an arc with $|J| \le \frac{1}{2}$ and $x, y \in J$. Then*

$$d\big((T^n)_i^{-1}(x), (T^n)_i^{-1}(y)\big) \le \lambda^{-n}\, d(x,y)$$

for all $1 \le i \le w^n$ and all $n \ge 1$. (Here $(T^n)_i^{-1}$ are the inverse branches of T^n on J.)

Proof. Consider the arc $J' \subset J$ with endpoints x, y. It is a subset of a semicircle in $\mathbb{S}^1$, so that $d(x,y) = |J'|$. Hence, $d\big((T^n)_i^{-1}(x), (T^n)_i^{-1}(y)\big) = |(T^n)_i^{-1}(J')| \le \lambda^{-n}|J'| = \lambda^{-n} d(x,y)$ by the observation preceding the lemma. □

2.2. **Transfer operator.** The transfer operator $\mathcal{L}_T : L^1(\mathfrak{m}) \to L^1(\mathfrak{m})$ is defined by the "duality equation":

$$\int_{\mathbb{S}^1} f \circ T \cdot g \, \mathrm{d}\mathfrak{m} = \int_{\mathbb{S}^1} f \cdot \mathcal{L}_T g \, \mathrm{d}\mathfrak{m} \qquad \forall\, f \in L^\infty(\mathfrak{m}), g \in L^1(\mathfrak{m}). \tag{6}$$

In particular, if ψ is the density of a probability measure ν, $\mathcal{L}\psi$ is the density of the push-forward measure $T_*\nu = \nu \circ T^{-1}$. In what follows, the T-subscript is omitted if there is no risk of misconception and $\mathcal{L}$ denotes the transfer operator associated with T. One can easily confirm that $\mathcal{L}u$ has the expression

$$\mathcal{L}u(x) = \sum_{y \in T^{-1}\{x\}} \frac{u(y)}{T'(y)}. \tag{7}$$

The missing details in the proof of the following lemma, which summarises the basic properties of $\mathcal{L}$, are left to the reader.

Lemma 2.3. *For transfer operator related to expanding mapping T,*

(1) $|\mathcal{L}u| \le \mathcal{L}|u|$ *for all* $u : \mathbb{S}^1 \to \mathbb{R}$,
(2) $0 \le u \le v \Longrightarrow 0 \le \mathcal{L}u \le \mathcal{L}v$ *for all* $u, v : \mathbb{S}^1 \to \mathbb{R}$,
(3) $\int_{\mathbb{S}^1} \mathcal{L}u \, \mathrm{d}\mathfrak{m} = \int_{\mathbb{S}^1} u \, \mathrm{d}\mathfrak{m}$ *for all* $u \in L^1(\mathfrak{m})$ *and*
(4) $\|\mathcal{L}u\|_{L^1(\mathfrak{m})} \le \|u\|_{L^1(\mathfrak{m})}$ *for all* $u \in L^1(\mathfrak{m})$.

Proof.
1. Use (7) and triangle inequality.
2. Use (7) and linearity of $\mathcal{L}$.
3. By Equation (6)

$$\int_{\mathbb{S}^1} \mathcal{L}u \, \mathrm{d}\mathfrak{m} = \int_{\mathbb{S}^1} 1 \cdot \mathcal{L}u \, \mathrm{d}\mathfrak{m} = \int_{\mathbb{S}^1} 1 \circ T \cdot u \, \mathrm{d}\mathfrak{m} = \int_{\mathbb{S}^1} u \, \mathrm{d}\mathfrak{m} \;\forall\, u \in L^2$$

4. Combine properties 1 and 3. □

One can also deduce expressions

$$\mathcal{L}u(x) = \sum_{y=1}^{w} \frac{u(T_i^{-1}(x))}{T'(T_i^{-1}(x))} \tag{8}$$

$$(\mathcal{L}u)'(x) = \sum_{i=1}^{w} \left(\frac{u'(T_i^{-1}(x))}{T'(T_i^{-1}(x))^2} - \frac{u(T_i^{-1}(x))}{T'(T_i^{-1}(x))^3} T''(T_i^{-1}(x)) \right) \tag{9}$$

in case $u \in C^1$, using the inverse branches of T.

3. Existence of a.c.i.m.

The existence of the a.c.i.m. is proved by a compactness argument similar to Arzelà–Ascoli theorem. More precisely, we are about to prove that C^1-norms of the pushforward densities remain uniformly bounded and bounded subsets of C^1 are compact in the space of Lipschitz functions endowed with its natural norm. It is worth mentioning that the strategy we use, that is, establishing invariance by taking averages along trajectories and using compactness to deduce that a limit exists, is widely used method to construct invariant measures (see e.g. the proof of Krylov–Bogolyubov Theorem [6, p. 151]).

3.1. **Boundedness.** In this section, we derive estimates for the transfer operator which C. Liverani motivated in [2]. The ultimate result we wish to prove is that for every fixed $\psi \in C^1$ there is $L > 0$ such that

$$\sup_{n\in\mathbb{N}} \|\mathcal{L}^n\psi\|_{C^1} = \sup_{n\in\mathbb{N}} \left(\|\mathcal{L}^n\psi\|_\infty + \|(\mathcal{L}^n\psi)'\|_\infty\right) \le L. \tag{10}$$

We begin by pointing out that the frequently used estimate

$$|(\mathcal{L}\psi)'| \le \frac{1}{\lambda}\mathcal{L}|\psi'| + \frac{\|T''\|_\infty}{\lambda^2}\mathcal{L}|\psi| \tag{11}$$

follows directly from (9). Using this estimate and the fact the $\mathcal{L}$ is a contraction in $L^1(\mathfrak{m})$ (property 4 in Lemma 2.3) we can control L^1-norms of the pushforwards. Let us prove that these are uniformly bounded with respect to n.

Lemma 3.1. *Let $\psi \in C^1$ be arbitrary. For every $n \in \mathbb{N}$ we have the following upper bound*

$$\|(\mathcal{L}^n\psi)'\|_{L^1(\mathfrak{m})} \le \frac{1}{\lambda^n}\|\psi'\|_{L^1(\mathfrak{m})} + \frac{\|T''\|_\infty}{\lambda(\lambda-1)}\|\psi\|_{L^1(\mathfrak{m})}. \tag{12}$$

Proof. The $n = 1$ case is a direct consequence of the contracting nature of $\mathcal{L}$ and (11):

$$\|(\mathcal{L}\psi)'\|_{L^1(\mathfrak{m})} \le \frac{1}{\lambda}\|\mathcal{L}|\psi'|\|_{L^1(\mathfrak{m})} + \frac{\|T''\|_\infty}{\lambda^2}\|\mathcal{L}|\psi|\|_{L^1(\mathfrak{m})} \le \frac{1}{\lambda}\|\psi'\|_{L^1(\mathfrak{m})} + \frac{\|T''\|_\infty}{\lambda(\lambda-1)}\|\psi\|_{L^1(\mathfrak{m})}.$$

Next we claim that

$$\|(\mathcal{L}^n\psi)'\|_{L^1(\mathfrak{m})} \le \frac{1}{\lambda^n}\|\psi'\|_{L^1(\mathfrak{m})} + \|T''\|_\infty\|\psi\|_{L^1(\mathfrak{m})}\sum_{i=2}^{n+1}\frac{1}{\lambda^i} \tag{13}$$

holds for any $n \in \mathbb{N}$. To this end, we proceed by induction. Suppose (13) holds for $n = k$. For $n = k + 1$ we have

$$\begin{aligned}
\|(\mathcal{L}^{k+1}\psi)'\|_{L^1(\mathfrak{m})} &\le \frac{1}{\lambda}\|\mathcal{L}|(\mathcal{L}^k\psi)'|\|_{L^1(\mathfrak{m})} + \frac{\|T''\|_\infty}{\lambda^2}\|\mathcal{L}|\mathcal{L}^k\psi|\|_{L^1(\mathfrak{m})} \\
&\le \frac{1}{\lambda}\|(\mathcal{L}^k\psi)'\|_{L^1(\mathfrak{m})} + \frac{\|T''\|_\infty}{\lambda^2}\|\psi\|_{L^1(\mathfrak{m})} \\
&\le \frac{1}{\lambda}\left(\frac{1}{\lambda^k}\|\psi'\|_{L^1(\mathfrak{m})} + \|T''\|_\infty\|\psi\|_{L^1(\mathfrak{m})}\sum_{i=2}^{k+1}\frac{1}{\lambda^i}\right) + \frac{\|T''\|_\infty}{\lambda^2}\|\psi\|_{L^1(\mathfrak{m})} \\
&\le \frac{1}{\lambda^{k+1}}\|\psi'\|_{L^1(\mathfrak{m})} + \|T''\|_\infty\|\psi\|_{L^1(\mathfrak{m})}\sum_{i=2}^{k+2}\frac{1}{\lambda^i}.
\end{aligned}$$

Thus, we conclude that (13) holds for any $n \in \mathbb{N}$. Finally, we observe that

$$\sum_{i=2}^{n+1}\frac{1}{\lambda^i} \le \frac{1}{\lambda(\lambda-1)} \ \forall\, n \in \mathbb{N}$$

which finishes the proof. $\square$

The constant $\|T''\|_\infty/(\lambda(\lambda-1))$ appearing in the left-hand side of (12) is so extensively used throughout this study that we denote it by the symbol Ω, i.e.,

$$\Omega := \frac{\|T''\|_\infty}{\lambda(\lambda-1)}. \tag{14}$$

Constant Ω in a way measures the "regularity" of the dynamics, that is, the linearity of the dynamics compared to the minimum magnitude of stretching.

Remark 3.1. *The contribution of the derivative of the initial funtion ψ to the upper bound in (12) becomes neglible after a sufficiently long time. This is a very general property of expanding dynamics; in the long term, sufficiently regular functions become even more smoothly distributed — up to a certain limit.*

The reason to examine L^1-norms is that in case of C^1 probability densities we can bound sup-norms by L^1-norms. More precisely, if ψ is differentiable density there is at least one point x_0 in which $\psi(x_0) = 1$. Otherwise it would not be possible to have $\int_{\mathbb{S}^1} \psi \, d\mathfrak{m} = 1$. Thus,

$$|\psi(x)| = |\int_{x_0}^{x} \psi'(t)dt + \psi(x_0)| \leq \int_{x_0}^{x} |\psi'(t)|dt + 1 \leq \|\psi'\|_{L^1(\mathfrak{m})} + 1$$

holds for every $x \in \mathbb{S}^1$ and therefore

$$\|\psi\|_\infty \leq \|\psi'\|_{L^1(\mathfrak{m})} + 1. \tag{15}$$

Remark that $\|\mathcal{L}^n v\|_\infty \leq \|v\|_\infty \|\mathcal{L}^n \mathbf{1}\|_\infty$ holds by the second part of lemma 2.3, where $\mathbf{1} = \chi_{\mathbb{S}^1} \in C^1$. We now have the tools to establish the desired boundedness result.

Theorem 3.1. *For every bounded ψ*

$$\sup_{n\in\mathbb{N}} \|\mathcal{L}^n \psi\|_\infty \leq (1+\Omega)\|\psi\|_\infty$$

Theorem 3.2. *For every continuously differentiable ψ*

$$\sup_{n\in\mathbb{N}} \|\mathcal{L}^n \psi\|_{C^1} \leq (1+\Omega)^2 \|\psi\|_{C^1}$$

Proof of the Theorem 3.1. Notice that $\mathcal{L}^n\mathbf{1}$ is C^1 for all $n \in \mathbb{N}$. Therefore, estimates (12) and (15) results in

$$\|\mathcal{L}^n \psi\|_\infty \leq \|\psi\|_\infty \|\mathcal{L}^n \mathbf{1}\|_\infty \leq \|\psi\|_\infty \left(\|(\mathcal{L}^n \mathbf{1})'\|_{L^1(\mathfrak{m})} + 1\right) \leq \|\psi\|_\infty \left(\frac{\|T''\|_\infty}{\lambda(\lambda-1)} + 1\right).$$

□

Proof of the Theorem 3.2. In order to achieve uniform bound, we first claim that

$$|(\mathcal{L}^n \psi)'| \leq \frac{1}{\lambda^n} \mathcal{L}^n |\psi'| + \|T''\|_\infty \sum_{i=2}^{n+1} \frac{\mathcal{L}^{n+2-i}|\psi|}{\lambda^i}.$$

By estimate (11) this holds when $n = 1$. For $n > 1$ we proceed by induction. Suppose the inequality holds for $n = k$. If $n = k + 1$ it holds that

$$\begin{aligned}
|(\mathcal{L}^{k+1}\psi)'| &\leq \frac{1}{\lambda}\mathcal{L}|(\mathcal{L}^k\psi)'| + \frac{\|T''\|_\infty}{\lambda^2}\mathcal{L}^{k+1}|\psi| \\
&\leq \frac{1}{\lambda}\mathcal{L}\left|\frac{1}{\lambda^k}\mathcal{L}^k|\psi'| + \|T''\|_\infty \sum_{i=2}^{k+1} \frac{\mathcal{L}^{k+2-i}|\psi|}{\lambda^i}\right| + \frac{\|T''\|_\infty}{\lambda^2}\mathcal{L}^{k+1}|\psi| \\
&\leq \frac{1}{\lambda^{k+1}}\mathcal{L}^{k+1}|\psi'| + \|T''\|_\infty \sum_{i=2}^{k+1} \frac{\mathcal{L}^{k+3-i}|\psi|}{\lambda^{i+1}} + \frac{\|T''\|_\infty}{\lambda^2}\mathcal{L}^{k+1}|\psi| \\
&= \frac{1}{\lambda^{k+1}}\mathcal{L}^{k+1}|\psi'| + \|T''\|_\infty \sum_{i=2}^{k+2} \frac{\mathcal{L}^{k+3-i}|\psi|}{\lambda^i}.
\end{aligned}$$

Therefore, by the induction principle, the estimate holds for any $n \in \mathbb{N}$. From this estimate it is easy to deduce the following bound

$$\|(\mathcal{L}^n\psi)'\|_\infty \leq (1+\Omega)(\|\psi'\|_\infty + \Omega\|\psi\|_\infty) \tag{16}$$

What we have proved is that

$$\begin{aligned}
\|\mathcal{L}^n\psi\|_{C^1} &\leq (1+\Omega)^2\|\psi\|_\infty + (1+\Omega)\|\psi'\|_\infty \\
&\leq (1+\Omega)^2\|\psi\|_\infty + (1+\Omega)^2\|\psi'\|_\infty \\
&= (1+\Omega)^2\|\psi\|_{C^1}
\end{aligned} \tag{17}$$

for all natural numbers n and continuously differentiable ψ. □

3.2. **Compactness.** We begin by proving the compactness result that guarantees the existence of a Lipschitz continuous invariant density.

Lemma 3.2. *Suppose that a sequence of C^1 functions $f_n : \mathbb{S}^1 \to \mathbb{R}$ is bounded, i.e., $\sup_n \|f_n\|_{C^1} \leq L < \infty$. Then it has a subsequence $(f_{n_k})_k$ which converges uniformly to a Lipschitz function f having Lipschitz constant* $\mathrm{Lip}(f) \leq L$.

Proof. The existence of the uniformly convergent subsequence is just an application of the Arzelà–Ascoli theorem: C^1–bounded sequence is, of course, equicontinuous and pointwise uniformly bounded so it satisfies the conditions of the theorem. Since the uniformly convergent exists, the proof is finished by noting that uniform limit of functions with uniformly bounded Lipschitz constants is Lipschitz continuous. □

The lemma can be generalised a little and re-interpreted as follows: bounded subsets of C^1 are relatively compact in the space of Lipschitz continuous functions endowed with its natural norm $\|\cdot\| = \|\cdot\|_\infty + \mathrm{Lip}(\cdot)$. With the aid of this lemma we are ready to prove the main theorem of the section.

Theorem 3.3 (The existence of the Lipschitz continuous invariant density)**.** *There is an invariant measure μ such that it is absolutely continuous w.r.t. Lebesgue measure and its density is Lipschitz continuous with its Lipschitz constant bounded by a quantity depending only on the dynamics of the system. More precisely, the density $\phi = \mathrm{d}\mu/\mathrm{d}\mathfrak{m}$ satisfies* $\mathrm{Lip}(\phi) \leq (1+\Omega)^2$.

Proof. Let ψ be C^1-density associated to absolutely continuous, not necessarily invariant, probability measure ν. By theorem 3.2, we know that densities $\mathcal{L}^n\psi$ are contained in

some closed ball in C^1. Now, define a sequence of probability measures μ_n by densities of the form

$$\psi_N = \frac{1}{N}\sum_{n=0}^{N-1} \mathcal{L}^n \psi. \tag{18}$$

Since the averages are contained in the closed ball, we know that there is a subsequence $(\psi)_{n_j}$ such that ψ_{n_j} converges uniformly to a Lipschitz density ϕ. Let μ be the probability measure defined by the limit density ϕ. In order to show it is invariant, let A be an arbitrary measurable set. By the bounded convergence theorem and trivial observation that

$$\int_{T^{-1}A} \psi_{n_j}\,\mathrm{d}\mathfrak{m} = \int_{\mathbb{S}^1} 1_{T^{-1}A}\,\psi_{n_j}\,\mathrm{d}\mathfrak{m} = \int_{\mathbb{S}^1} 1_A \circ T\,\psi_{n_j}\,\mathrm{d}\mathfrak{m} = \int_{\mathbb{S}^1} 1_A\,\mathcal{L}\psi_{n_j}\,\mathrm{d}\mathfrak{m} = \int_A \mathcal{L}\psi_{n_j}\,\mathrm{d}\mathfrak{m},$$

we have the following equality

$$\begin{aligned}
\mu(T^{-1}A) - \mu(A) &= \int_{T^{-1}A} \lim_{j\to\infty} \psi_{n_j}\,\mathrm{d}\mathfrak{m} - \int_A \lim_{j\to\infty} \psi_{n_j}\,\mathrm{d}\mathfrak{m} \\
&= \lim_{j\to\infty} \int_{T^{-1}A} \psi_{n_j}\,\mathrm{d}\mathfrak{m} - \lim_{j\to\infty} \int_A \psi_{n_j}\,\mathrm{d}\mathfrak{m} \\
&= \lim_{j\to\infty} \int_A (\mathcal{L}\psi_{n_j} - \psi_{n_j})\,\mathrm{d}\mathfrak{m}.
\end{aligned}$$

Now the proof of invariance is accomplished by the following estimate:

$$\begin{aligned}
\left|\int_A (\mathcal{L}\psi_{n_j} - \psi_{n_j})\,\mathrm{d}\mathfrak{m}\right| &= \left|\int_A \frac{\mathcal{L}^{n_j}\psi}{n_j} - \frac{\psi}{n_j}\mathrm{d}\mathfrak{m}\right| \\
&\leq \frac{1}{n_j}(\|\mathcal{L}^{n_j}\psi\|_{L^1} + \|\psi\|_{L^1}) \\
&\leq \frac{2\|\psi\|_{L^1}}{n_j} \longrightarrow 0.
\end{aligned}$$

An upper bound for Lipschitz constant is obtained by estimates on the Lipschitz constants of ψ_{n_j}:

$$\begin{aligned}
\mathrm{Lip}(\phi) &\leq \sup_{j\in\mathbb{N}} \mathrm{Lip}(\psi_{n_j}) \\
&\leq \sup_{j\in\mathbb{N}} \|\psi_{n_j}\|_{C^1} \\
&\leq (1+\Omega)^2 \|\psi\|_{C^1}
\end{aligned}$$

Additionally, since ψ was arbitrary, we may pick $\psi \equiv 1$ to obtain

$$\mathrm{Lip}(\phi) \leq (1+\Omega)^2. \tag{19}$$

□

Remark that this is just an existence theorem, not uniqueness; we have showed that *every* C^1 density ψ results in *at least one* invariant density. Nevertheless, uniqueness will follow from the convergence of an arbitrary density towards the invariant density.

4. Exponential decay of correlation coefficients

Recall that the mixing property stated in terms of decay of *correlation coefficients* reads

$$\lim_{n\to\infty}\int f\circ T^n\cdot g\,\mathrm{d}\mu=\int f\,\mathrm{d}\mu\int g\,\mathrm{d}\mu \qquad \forall\, f,g\in L^2(\mu). \tag{20}$$

In previous section, we proved that μ is absolutely continuous with respect to the Lebesgue measure. Denoting the density by ϕ, mixing is equivalent to

$$\lim_{n\to\infty}\int f\circ T^n\cdot g\phi\,\mathrm{d}\mathfrak{m}=\int f\phi\,\mathrm{d}\mathfrak{m}\int g\phi\,\mathrm{d}\mathfrak{m} \qquad \forall\, f,g\in L^2(\mu).$$

As motivated at the beginning of the paper, we are willing to restrict ourselves to Hölder continuous g in order to obtain concrete estimates on the error terms

$$\left|\int f\circ T^n\cdot g\phi\,\mathrm{d}\mathfrak{m}-\int f\phi\,\mathrm{d}\mathfrak{m}\int g\phi\,\mathrm{d}\mathfrak{m}\right|.$$

To this end, by adding and subtracting $2\int f\phi\,\mathrm{d}\mathfrak{m}\|g\|_\infty$, and by normalizing, we get

$$\left|\int g\,\mathrm{d}\mu+2\|g\|_\infty\right|\cdot\left|\int f\circ T^n\cdot\frac{\phi(g+2\|g\|_\infty)}{\int g\,\mathrm{d}\mu+2\|g\|_\infty}\,\mathrm{d}\mathfrak{m}-\int f\phi\,\mathrm{d}\mathfrak{m}\right|.$$

Denoting

$$\psi=\frac{\phi(g+2\|g\|_\infty)}{\int g\,\mathrm{d}\mu+2\|g\|_\infty},$$

we can reduce the decay of correlation coefficients to $L^1(\mathfrak{m})$–convergence of probability densities towards the invariant density:

$$\left|\int f\circ T^n\cdot g\phi\,\mathrm{d}\mathfrak{m}-\int f\phi\,\mathrm{d}\mathfrak{m}\int g\phi\,\mathrm{d}\mathfrak{m}\right|\le 3\|g\|_\infty\|f\|_\infty\|\mathcal{L}^n\psi-\phi\|_{L^1(\mathfrak{m})}.$$

Therefore, it is sufficient to prove the inequality

$$\|\mathcal{L}^n\psi-\phi\|_{L^1(\mathfrak{m})}\le D\theta^n \qquad \forall\,\psi\in C^\alpha \tag{21}$$

with $D>0$, $\theta\in(0,1)$, in order to achieve exponentially decaying estimates. This will be proved after we have developed some tools needed to construct the coupling argument.

In what follows, we use the notation

$$H_\alpha(f)=\sup_{x,y\in X}\frac{|f(x)-f(y)|}{d(x,y)^\alpha}$$

for the Hölder coefficient of a function f and denote the class of Hölder continuous functions of order α by C^α. That is,

$$C^\alpha=\{f:X\to\mathbb{R}:H_\alpha(f)<\infty\}$$

4.1. **Preliminary results.** The following *distortion bound* is central for understanding the structure of push-forward densities $\mathcal{L}^n\psi$. The core idea of the proof is basically the same as Lai-Sang Young's in [3, p. 33-34].

Lemma 4.1. *Let $n\in\mathbb{N}$ be arbitrary. For any $x,y\in\mathbb{S}^1$,*

$$e^{-\Omega d(x,y)}\le\frac{(T^n)'((T^n)_i^{-1}x)}{(T^n)'((T^n)_i^{-1}y)}\le e^{\Omega d(x,y)}, \tag{22}$$

in which $\Omega=\|T''\|_\infty/(\lambda(\lambda-1))$ is independent of n. *Here $(T^n)_i^{-1}$ is the ith branch of the inverse of T^n on a given arc $J\subset\mathbb{S}^1$ of length $|J|\le\frac{1}{2}$ containing both x and y.*

Proof. Let J be as in the statement of the lemma. Lemma 2.2 implies that, for an arbitrary $k \geq 1$, the inverse branches of T^k on J are well defined. For brevity, let x_{-k} and y_{-k} denote the preimages of x and y, respectively, along the same branch.

Since the logarithm is λ^{-1}–Lipschitz in $[\lambda, \|T'\|_\infty]$, we can estimate

$$\begin{aligned}
\log \frac{(T^n)'(x_{-n})}{(T^n)'(y_{-n})} &\leq |\log((T^n)'(x_{-n})) - \log((T^n)'(y_{-n}))| \\
&\leq \sum_{i=1}^{n-1} |\log(T'(T^i x_{-n})) - \log(T'(T^i y_{-n}))| \\
&\leq \sum_{i=1}^{n-1} \frac{1}{\lambda} \|T''\|_\infty d(x_{-n+i}, y_{-n+i}) \\
&\leq \sum_{i=1}^{n-1} \frac{1}{\lambda} \|T''\|_\infty \lambda^{-n+i} d(x, y) \\
&\leq \frac{\|T''\|_\infty}{\lambda(\lambda-1)} d(x, y) = \Omega d(x, y).
\end{aligned}$$

A similar estimate is obtained by interchanging x and y, which proves the claim. $\square$

This lemma is immediately used in the following theorem which is the cornerstone of our coupling argument. The trick is that sometimes, particularly in our case, it is more convenient to work with logarithms of functions instead of functions themselves.

Theorem 4.1. *Suppose ψ is a strictly positive probability density and that $\log \psi \in C^\alpha$. Then $\mathcal{L}^n \psi$ has the same properties for every $n \in \mathbb{N}$ and*

$$H_\alpha(\log \mathcal{L}^n \psi) \leq \frac{H_\alpha(\log \psi)}{\lambda^{\alpha n}} + \Omega.$$

Proof. Let $J \subset \mathbb{S}^1$ be an interval with $|J| \leq \frac{1}{2}$. Given an initial probability density ψ, we introduce the notation

$$\psi_{n,i}(x) = \frac{\psi((T^n)_i^{-1} x)}{(T^n)'((T^n)_i^{-1} x)}\ , \qquad x \in J,$$

where $(T^n)_i^{-1}$ denotes the ith inverse branch of T on J. Then

$$\mathcal{L}^n \psi(x) = \sum_{i=1}^{w^n} \psi_{n,i}(x).$$

The number w has been introduced in Section 2.

Next, let $x, y \in \mathbb{S}^1$ be arbitrary. Without loss of generality, we may assume both points belong to J. Therefore,

$$\begin{aligned}
\left| \log \frac{\psi_{n,i}(x)}{\psi_{n,i}(y)} \right| &\leq \left| \log \frac{\psi((T^n)_i^{-1} x)}{\psi((T^n)_i^{-1} y)} \right| + \left| \log \frac{(T^n)'((T^n)_i^{-1} y)}{(T^n)'((T^n)_i^{-1} x)} \right| \\
&\leq H_\alpha(\log \psi)\, d((T^n)_i^{-1} x, (T^n)_i^{-1} y)^\alpha + \Omega d(x, y) \\
&\leq \lambda^{-\alpha n} H_\alpha(\log \psi)\, d(x, y)^\alpha + \Omega d(x, y)^\alpha.
\end{aligned}$$

For brevity, denote the right side of the last bound by R. Then

$$e^{-R} \leq \frac{\psi_{n,i}(x)}{\psi_{n,i}(y)} \leq e^R, \qquad i.e. \qquad e^{-R} \psi_{n,i}(y) \leq \psi_{n,i}(x) \leq e^R \psi_{n,i}(y).$$

Summing over i, we get

$$e^{-R}\mathcal{L}^n\psi(y) \leq \mathcal{L}^n\psi(x) \leq e^R\mathcal{L}^n\psi(y), \qquad i.e. \qquad e^{-R} \leq \frac{\mathcal{L}^n\psi(x)}{\mathcal{L}^n\psi(y)} \leq e^R.$$

Taking logarithms, this yields

$$\left|\log \frac{\mathcal{L}^n\psi(x)}{\mathcal{L}^n\psi(y)}\right| \leq (\lambda^\alpha)^{-n} H_\alpha(\log\psi)\, d(x,y)^\alpha + \Omega d(x,y)^\alpha,$$

which is the desired bound. $\square$

We finish this section by demonstrating how estimates on Hölder constants of logarithms can be used to obtain information about approppriate *probability* densities.

As mentioned earlier, for every probability density ψ, there is a point x_0 such that $\psi(x_0) = 1$. Therefore,

$$|\log(\psi(x))| = |\log(\psi(x)) - \log(\psi(x_0))| \leq H_\alpha(\log(\psi))d(x,x_0)^\alpha \leq H_\alpha(\log(\psi)),$$

i.e.

$$\exp(-H_\alpha(\log(\psi))) \leq \psi(x) \leq \exp(H_\alpha(\log(\psi))), \tag{23}$$

holds for every $x \in \mathbb{S}^1$. Furthermore, for every $x, y \in \mathbb{S}^1$

$$d(x,y)^\alpha H_\alpha(\log\psi) \geq |\log(\psi(x)) - \log(\psi(y))| = |\int_{\psi(y)}^{\psi(x)} \frac{dt}{t}| \geq \frac{1}{\|\psi\|_\infty}|\psi(x) - \psi(y)|$$

from which it follows that

$$H_\alpha(\psi) \leq H_\alpha(\log\psi)\|\psi\|_\infty \leq H_\alpha(\log\psi)\exp\left(H_\alpha(\log\psi)\right). \tag{24}$$

Therefore, Hölder continuity of the logarithm implies Hoelder continuity of the density itself. Additionally, it results in upper and lower bounds on the density.

4.2. **Coupling.** We begin by proving the following theorem which summarises the relevant consequences of the results of the previous subsection. To this end, let us introduce the following notation:

$$\mathcal{H}_D = \{\psi : X \to \mathbb{R} : \psi > 0, H_\alpha(\log\psi) \leq D\}.$$

$\alpha \in (0,1]$ is fixed and $D \in \mathbb{R}_+$.

Theorem 4.2.

(1) *For all $B > 0$ there is $N = N(B) \in \mathbb{N}$ such that for all $\psi \in \mathcal{H}_B$ and $n > N$ $\mathcal{L}^n\psi \in \mathcal{H}_{\Omega+1}$*
(2) *There is $a > 0$ such that $\psi \geq 2a > 0$ holds for every $\psi \in \mathcal{H}_{\Omega+1}$.*
(3) *There is K such that $\tilde{\psi} := (\psi - a)/(1-a) \in \mathcal{H}_K$ for all $\psi \in \mathcal{H}_{\Omega+1}$. The constant a is the one from the previous item.*

Proof.

(1) By theorem 4.1, this holds when

$$\frac{B}{\lambda^{\alpha N}} + \Omega < \Omega + 1 \text{ i.e. } N > \frac{\log(B)}{\alpha\log(\lambda)}.$$

Thus, $N(B)$ can be chosen to be the integer part of $\log(B)/(\alpha\log(\lambda)) + 1$.

(2) By equation (23),

$$\psi(x) \geq \exp\left(-H_\alpha(\log\psi)\right) \geq \exp\left(-(\Omega+1)\right) =: 2a$$

holds for every $\psi \in \mathcal{H}_{\Omega+1}$.

(3) By equations (23) and (24),

$$\begin{aligned}
&|\log(\frac{\psi(x)-a}{1-a})-\log(\frac{\psi(y)-a}{1-a})| \\
&\leq |\psi(x)-\psi(y)| \sup_x \frac{1}{\psi(x)-a} \\
&\leq \frac{1}{a} H_\alpha(\log\psi)\exp\left(\frac{H_\alpha(\log\psi)}{2^\alpha}\right) d(x,y)^\alpha \\
&\leq 2(\Omega+1)\exp\left(2(\Omega+1)\right) d(x,y)^\alpha \\
&\leq \exp\left(4(\Omega+1)\right) d(x,y)^\alpha \\
&=: K d(x,y)^\alpha
\end{aligned}$$

holds for all $\psi \in \mathcal{H}_{\Omega+1}$. Thus, $\tilde{\psi} \in \mathcal{H}_K$ for all $\psi \in \mathcal{H}_{\Omega+1}$. In particular,

$$N(K) = \frac{4(\Omega+1)}{\alpha\log(\lambda)}.$$

□

Now we are ready to proceed to the coupling argument itself. In what follows, let α and T be fixed and K and a be the constants from the previous theorem. Let us prove the exponential convergence first for functions in $\mathcal{H}_K$.

Let ψ_1 and ψ_2 be arbitrary densities from $\mathcal{H}_K$. After time $N = N(K)$ calculated earlier the pushforwards of these densities lie in $\mathcal{H}_{\Omega+1}$. Thus, for $n > N$

$$\mathcal{L}^n\psi_i = a + (1-a)\tilde{\psi}_i$$

for some unique $\tilde{\psi}_i \in \mathcal{H}_K$. Using this equality, $\mathcal{L}$ being linear contraction in $L^1(\mathfrak{m})$ yields

$$\|\mathcal{L}^n\psi_1 - \mathcal{L}^n\psi_2\|_{L^1(\mathfrak{m})} \leq (1-a)\|\tilde{\psi}_1 - \tilde{\psi}_2\|_{L^1(\mathfrak{m})}.$$

Remark that $L^1(\mathfrak{m})$-norms of functions in $\mathcal{H}_B$ are uniformly bounded by $\exp(B/2^\alpha)$ for every $B > 0$. Now we can apply the same argument for $\tilde{\psi}_1$ and $\tilde{\psi}_2$ which belong to $\mathcal{H}_K$ as well. Therefore, for every $n > 2N$ the inequality

$$\|\mathcal{L}^n\psi_1 - \mathcal{L}^n\psi_2\|_{L^1(\mathfrak{m})} \leq (1-a)^2\|\tilde{\psi}_1^1 - \tilde{\psi}_2^1\|_{L^1(\mathfrak{m})}.$$

holds for some $\tilde{\psi}_i^1 \in \mathcal{H}_K$. Moreover, by simple induction argument we deduce that

$$n \in [kN, (k+1)N] \implies \|\mathcal{L}^n\psi_1 - \mathcal{L}^n\psi_2\|_{L^1(\mathfrak{m})} \leq (1-a)^k\|\tilde{\psi}_1^k - \tilde{\psi}_2^k\|_{L^1(\mathfrak{m})}$$

holds for every $k \in \mathbb{N}$ with some unique $\tilde{\psi}_i^k \in \mathcal{H}_K$. On the other hand, $n/N \leq (k+1)$ yields

$$(1-a)^k \leq (1-a)^{n/N-1}.$$

Combining these results and uniform bound of norms yields

$$n \in [kN, (k+1)N] \implies \|\mathcal{L}^n\psi_1 - \mathcal{L}^n\psi_2\|_{L^1(\mathfrak{m})} \leq (1-a)^{n/N-1} 2\exp(K/2^\alpha)$$

Since this holds for every $k \in \mathbb{N}$, we have proved the exponential convergence

$$\|\mathcal{L}^n\psi_1 - \mathcal{L}^n\psi_2\|_{L^1(\mathfrak{m})} \leq D\theta^{\alpha n} \tag{25}$$

with parameters

$$D = \frac{2\exp(K2^{-\alpha})}{1-a} \quad \text{and} \quad \theta = (1-a)^{\frac{1}{\alpha N}} \in (0,1). \tag{26}$$

Using parameters of the system (and the fact that a must be smaller than $1/2$):

$$D \leq \frac{2\exp(\exp(4(\Omega+1)))}{1-1/2\exp(-(\Omega+1))} \leq 4\exp(\exp(4(\Omega+1))) \tag{27}$$

$$\theta = (1-\frac{1}{2}\exp(-(\Omega+1))^{\log(\lambda)/(4(\Omega+1))} \leq (1-\exp(-3(\Omega+1)))^{\log(\lambda)/(4(\Omega+1))}. \tag{28}$$

To justify the use of the term "coupling" earlier, let us interpret our results in a more probabilistic fashion. First, take two initial densities ψ^1 and ψ^2. Consider random variable X and Y taking values in $\mathbb{S}^1$ according to ψ^1 and ψ^2, respectively. Then, for any $n \geq 1$, $X_n = T^n(X)$ and $Y_n = T^n(Y)$ are distributed according to $\mathcal{L}^n\psi^1$ and $\mathcal{L}^n\psi^2$, respectively. We construct a coupling $Z_n = (X'_n, Y'_n)$ of X_n and Y_n, in the probabilistic sense of the word, for each $n \geq 1$. To that end, notice from our construction that there exists a sequence of times $0 = n_0 < n_1 < n_2 < \cdots$ with the properties that, for any $k \geq 0$ and $n_k \leq n < n_{k+1}$,

$$\mathcal{L}^n\psi^i = (1-a)^k\psi^i_n + (1-(1-a)^k)\rho_n, \tag{29}$$

where ψ^i_n and ρ_n are probability densities, and ρ_n does not depend on ψ^i at all. A good definition of Z_n is obtained as follows. Toss a weighted coin with the probability of heads being $1-(1-a)^k$. In the case of heads, we set $X'_n = Y'_n = R_n$ where R_n has the density ρ_n on $\mathbb{S}^1$. In the case of tails, we let X'_n and Y'_n be distributed according to ψ^1_n and ψ^2_n, respectively, independently of each other. It is easy to check that the distributions of X'_n and Y'_n are the same as those of X_n and Y_n. Thus, Z_n is indeed a coupling of X_n and Y_n.

A widely known result in probability theory states that the following coupling inequality holds for *an arbitrary* coupling:

$$\|\mathcal{L}^n\psi^1 - \mathcal{L}^n\psi^2\|_{L^1(\mathfrak{m})} \leq 2\mathbb{P}(X'_n \neq Y'_n).$$

Using the specific coupling introduced above, we have, for $n_k \leq n < n_{k+1}$, that $X'_n = Y'_n = R_n$ with probability $1-(1-a)^k$, which implies $\mathbb{P}(X'_n \neq Y'_n) \leq (1-a)^k$. This leads to the very same exponentially decaying estimate as in the previous section. (Of course, we could have just used (29). The point here is that (29) was used only in the construction of the coupling Z_n; had we been given Z_n directly, we could have derived the exponential bound again without 29. In any case, we have demonstrated that the work in the previous sections is genuinely related to the idea of coupling as understood by probabilists.)

A more comprehensive, standard introduction to the coupling techniques is Torgny Lindvall's [4]. It may also be instructive to compare our coupling argument to the one used in the classical proof of the exponential convergence of irreducible and aperiodic finite Markov chains towards a unique equilibrium distribution [5].

Next, we extend our results to the general Hölder densities. This results in the desired bound for the correlation coefficients on page 8. Although this is a straighforward consequence of Theorem 4.3, we establish the exact bound in Theorem 4.4 since this is the main result of the section.

Theorem 4.3 (Exponential convergence of densities)**.**
There exist constants $\tilde{D} \in \mathbb{R}$ and $\theta \in (0,1)$ such that

$$\|\mathcal{L}^n\psi - \phi\|_{L^1(\mathfrak{m})} \leq \tilde{D}(1+H_\alpha(\psi))\theta^{\alpha n}$$

holds for every $n \in \mathbb{N}$, for every $\psi \in C^\alpha$, for every (fixed) $\alpha \in (0,1)$.

Proof. First, remark that $H_\alpha(\log(\mathcal{L}^n\phi)) \leq 1+\Omega$ for sufficiently large n, by Corollary A.1 and Theorem 4.1. Thus, the invariance of ϕ results in $H_\alpha(\log(\phi)) \leq 1+\Omega \leq K$ and we can apply previous theorems to ϕ as well.

We can split $\psi = 2(\psi+1)/2 - 1$. Thus,

$$\begin{aligned}
\|\mathcal{L}^n\psi - \phi\|_{L^1(\mathfrak{m})} &= \|2\mathcal{L}^n\frac{\psi+1}{2} - \mathcal{L}^n\mathbf{1} - \phi\|_{L^1(\mathfrak{m})} \\
&= \|2\mathcal{L}^n\frac{\psi+1}{2} - 2\phi - (\mathcal{L}^n\mathbf{1} - \phi)\|_{L^1(\mathfrak{m})} \\
&\leq 2\|\mathcal{L}^n\frac{\psi+1}{2} - \phi\|_{L^1(\mathfrak{m})} + \|\mathcal{L}^n\mathbf{1} - \phi\|_{L^1(\mathfrak{m})} \\
&\leq 2\|\mathcal{L}^n\frac{\psi+1}{2} - \phi\|_{L^1(\mathfrak{m})} + D\theta^{-\alpha n},
\end{aligned}$$

where D and θ are from previous results (Since $H_\alpha(\mathbf{1}) = 0$). Additionally, $(\psi+1)/2 \in \mathcal{H}_{H_\alpha(\psi)}$ since

$$|\log(\frac{\psi(x)+1}{2}) - \log(\frac{\psi(y)+1}{2})| \leq \sup_x(\frac{2}{\psi(x)+1})|\frac{\psi(x)+1}{2} - \frac{\psi(y)+1}{2}| \leq H_\alpha(\psi)d(x,y)^\alpha.$$

That is,

$$H_\alpha(\log((\psi+1)/2)) \leq H_\alpha(\psi).$$

If $H_\alpha(\psi) \leq K$ (K from theorem 4.2), we have already proved exponential convergence of the form:

$$\|\mathcal{L}^n\psi - \phi\|_{L^1(\mathfrak{m})} \leq 3D\theta^{\alpha n}.$$

On the other hand, if $H_\alpha(\psi) > K$, $(\psi+1)/2 \in \mathcal{H}_{1+\Omega}$ after time $N(H_\alpha(\psi)) =: N_0$ and by previous results

$$\|\mathcal{L}^n\frac{\psi+1}{2} - \phi\|_{L^1} \leq D\theta^{\alpha(n-N_0)} \qquad \forall\ n > N_0.$$

If $n \leq N_0$,

$$\|\mathcal{L}^n\frac{\psi+1}{2} - \phi\|_{L^1} \leq \|(\psi+1)/2\|_{L^1} + \|\phi\|_{L^1} \leq 2\theta^{\alpha n}\theta^{-\alpha N_0}$$

Therefore, for every $n \in \mathbb{N}$ it holds that

$$\|\mathcal{L}^n\frac{\psi+1}{2} - \phi\|_{L^1} \leq \max\{D\theta^{-\alpha N_0}, 2\theta^{-\alpha N_0}\}\theta^{\alpha n} = D\theta^{-\alpha N_0}\theta^{\alpha n},$$

if D is chosen to be $4\exp(\exp(4(\Omega+1)))$ as previously. Thus,

$$\|\mathcal{L}^n\psi - \phi\|_{L^1} \leq D(2\theta^{-\alpha N_0} + 1)\theta^{\alpha n} \qquad \forall\ n \in \mathbb{N}.$$

Combining the results with $\mathcal{H}_{H_\alpha(\psi)} > K$ and $\mathcal{H}_{H_\alpha(\psi)} \leq K$ results in

$$\|\mathcal{L}^n\psi - \phi\|_{L^1} \leq \max\{3, 2\theta^{-\alpha N_0} + 1\}D\theta^{\alpha n} = (1 + 2\theta^{-\alpha N_0})D\theta^{\alpha n}.$$

Plug in the appropriate values[2]:

$$\begin{aligned}
\|\mathcal{L}^n\psi - \phi\|_{L^1} &\leq (1 + 2\cdot(1 - e^{-3(\Omega+1)})^{-(\log(H_\alpha(\psi))/4(\Omega+1))})4\exp(\exp(4(\Omega+1)))\theta^{\alpha n} \\
&\leq (1 + 2\cdot 2^{(\log(H_\alpha(\psi))/4(\Omega+1))})4\exp(\exp(4(\Omega+1)))\theta^{\alpha n} \\
&\leq (1 + 2\cdot H_\alpha(\psi))4\exp(\exp(4(\Omega+1)))\theta^{\alpha n} \\
&\leq 8\exp(\exp(4(\Omega+1)))(1 + H_\alpha(\psi))\theta^{\alpha n}.
\end{aligned}$$

Thus, we have proved the exponential estimate for all $\psi \in C^\alpha$ with

$$\begin{aligned}
\tilde{D} &= 8\exp(\exp(4(\Omega+1))) \\
\theta &= (1 - \exp(-3(\Omega+1))^{(\log(\lambda)/(4(\Omega+1)))}.
\end{aligned}$$

□

[2] There is no loss of generality in assuming that $H_\alpha(\log(\psi)) > 1$ since if this not the case, $\theta^{-\alpha N_0} < 1$ and the bound we obtained holds as well.

Finally, we combine the relevant results of this section and establish the exact exponential bound for the correlation coefficients.

Theorem 4.4 (Exponential decay of correlation coefficients). *For bounded f and Hölder continuous g,*

$$\left|\int_{\mathbb{S}^1} f\circ T^n g\,\mathrm{d}\mu - \int_{\mathbb{S}^1} f\,\mathrm{d}\mu \int_{\mathbb{S}^1} g\,\mathrm{d}\mu\right| \leq C\|f\|_\infty(\|g\|_\infty + H_\alpha(g))\theta^{\alpha n}, \tag{30}$$

where

$$C = 12(2+\Omega)^2\tilde{D}$$

and $\tilde{D}$ and θ are from previous theorem. In particular, $\tilde{D}$, θ and C are independent of f, g and α.

Proof. Considerations on page 8 results in

$$\left|\int_{\mathbb{S}^1} f\circ T^n g\mathrm{d}\mu - \int_{\mathbb{S}^1} f\,\mathrm{d}\mu \int_{\mathbb{S}^1} g\,\mathrm{d}\mu\right| \leq 3\|g\|_\infty\|f\|_\infty \left\|\mathcal{L}^n\left(\frac{\phi(g+2\|g\|_\infty)}{\mu(g)+2\|g\|_\infty}\right) - \phi\right\|_{L^1(\mathfrak{m})}$$

where $g \neq \bar{0}$ and ϕ is the invariant density. Due to Theorem 3.3, it is evident that

$$\begin{aligned}
&H_\alpha\left(\frac{\phi(g+2\|g\|_\infty)}{\mu(g)+2\|g\|_\infty}\right)\\
&\leq \frac{H_\alpha(\phi(g+2\|g\|_\infty))}{\mu(g)+2\|g\|_\infty}\\
&\leq \frac{\|\phi\|_\infty}{\|g\|_\infty}H_\alpha(g) + 3H_\alpha(\phi)\\
&\leq \frac{\mathrm{Lip}(\phi)+1}{\|g\|_\infty}H_\alpha(g) + 3\mathrm{Lip}(\phi) \leq (\frac{H_\alpha(g)}{\|g\|_\infty}+3)(2+\Omega)^2
\end{aligned}$$

and we can therefore use theorem 4.3 to bound the difference:

$$\begin{aligned}
&\left|\int_{\mathbb{S}^1} f\circ T^n g\mathrm{d}\mu - \int_{\mathbb{S}^1} f\,\mathrm{d}\mu \int_{\mathbb{S}^1} g\,\mathrm{d}\mu\right|\\
&\leq 3\|g\|_\infty\|f\|_\infty\tilde{D}\left(1+(\frac{H_\alpha(g)}{\|g\|_\infty}+3)(2+\Omega)^2)\right)\theta^{\alpha n}\\
&\leq 12\|f\|_\infty\tilde{D}(2+\Omega)^2\left(\|g\|_\infty + H_\alpha(g)\right)\theta^{\alpha n},
\end{aligned}$$

where $\tilde{D}$ and θ are as earlier (see the end of Theorem 4.3). □

5. Mixing

Finally, we are going to prove the mixing property of the dynamics, that is,

$$\int_{\mathbb{S}^1} f\circ T^n g\,\mathrm{d}\mu \longrightarrow \int_{\mathbb{S}^1} f\,\mathrm{d}\mu \int_{\mathbb{S}^1} g\,\mathrm{d}\mu \qquad \forall\ f,g \in L^2 \tag{31}$$

by generalising the known result for Hölder-functions (α being fixed) step by step:

1. (31) holds for bounded f and $g \in C^\alpha$.
2. for bounded f and $g \in C^0$.
3. for bounded f and simple g.
4. for bounded, measurable f and g.
5. for $f, g \in L^2$.

In what follows, asumming a function f is integrable we use extensively the following abbreviated notation:

$$\mu(f) = \int_{\mathbb{S}^1} f \, \mathrm{d}\mu.$$

Step 1: We have already proved this in Theorem 4.4.

Step 2: C^α is obviously a subalgebra of C^0 which contains constants. Additionally, the identity map (which separates points) belongs to C^α and therefore, by Stone–Weierstrass theorem, C^α is dense in C^0. Thus, for every $f, g \in C^0$ and $\epsilon > 0$ there are $\tilde{f}, \tilde{g} \in C^\alpha$ such that $\|f - \tilde{f}\|_\infty < \epsilon$ and $\|g - \tilde{g}\|_\infty < \epsilon$. Therefore

$$\begin{aligned}
&|\mu(f \circ T^n g) - \mu(f)\mu(g)| \\
&= |\mu(f \circ T^n g) - \mu(f \circ T^n \tilde{g}) + \mu(f \circ T^n \tilde{g}) - \mu(\tilde{f} \circ T^n \tilde{g}) + \mu(\tilde{f} \circ T^n \tilde{g}) \\
&\quad - \mu(f)\mu(g) + \mu(f)\mu(\tilde{g}) - \mu(f)\mu(\tilde{g}) + \mu(\tilde{f})\mu(\tilde{g}) - \mu(\tilde{f})\mu(\tilde{g})| \\
&\leq \mu(|f \circ T^n||g - \tilde{g}|) + \mu(|f - \tilde{f}| \circ T^n|\tilde{g}|) + \mu(|f|)\mu(|g - \tilde{g}|) \\
&\quad + \mu(|\tilde{g}|)\mu(|f - \tilde{f}|) + |\mu(\tilde{f} \circ T^n \tilde{g}) - \mu(\tilde{f})\mu(\tilde{g})| \\
&\leq 2\epsilon(\|\tilde{f}\|_{L^1(\mu)} + \|\tilde{g}\|_{L^1(\mu)}) + |\mu(\tilde{f} \circ T^n \tilde{g}) - \mu(\tilde{f})\mu(\tilde{g})| \\
&< 2\epsilon(\|\tilde{f}\|_{L^1(\mu)} + \|\tilde{g}\|_{L^1(\mu)} + 1)
\end{aligned}$$

when n is large enough. This proves step 2.

Step 3: As a consequence of Lusin's theorem, for every simple f, g there are continuous $\tilde{f}, \tilde{g}$ such that $\|f - \tilde{f}\|_{L^1(\mathfrak{m})} < \epsilon$ and $\|g - \tilde{g}\|_{L^1(\mathfrak{m})} < \epsilon$. Therefore, using the fact that the claim holds for continuous functions and $\|h\|_{L^1(\mathfrak{m})} \leq \|h\|_\infty \ \forall \ h$:

$$\begin{aligned}
&|\mu(f \circ T^n g) - \mu(f)\mu(g)| \\
&\leq \mu(|f \circ T^n||g - \tilde{g}|) + \mu(|f - \tilde{f}| \circ T^n|\tilde{g}|) + \mu(|f|)\mu(|g - \tilde{g}|) \\
&\quad + \mu(|\tilde{g}|)\mu(|f - \tilde{f}|) + |\mu(\tilde{f} \circ T^n \tilde{g}) - \mu(\tilde{f})\mu(\tilde{g})| \\
&\leq 2\epsilon(\|\tilde{f}\|_\infty + \|\tilde{g}\|_\infty) + |\mu(\tilde{f} \circ T^n \tilde{g}) - \mu(\tilde{f})\mu(\tilde{g})| \\
&< 2\epsilon(\|\tilde{f}\|_\infty + \|\tilde{g}\|_\infty + 1)
\end{aligned}$$

when n is large enough.

Step 4: Similar to step 2. In compact measurable spaces, bounded measurable functions can be uniformly approximated by simple functions.

Step 5: Remark, that bounded, measurable, functions are dense in L^2 by simple truncation argument. Therefore, for every $f, g \in L^2$ and $\epsilon > 0$ there are bounded, measurable, $\tilde{f}, \tilde{g}$ such that $\|\tilde{f} - f\|_{L^2(\mathfrak{m})} < \epsilon$ and $\|\tilde{f} - f\|_{L^2(\mathfrak{m})} < \epsilon$. Consequently,

$$\begin{aligned}
&|\mu(f \circ T^n g) - \mu(f)\mu(g)| \\
&\leq \mu(|f \circ T^n||g - \tilde{g}|) + \mu(|f - \tilde{f}| \circ T^n|\tilde{g}|) + \mu(|f|)\mu(|g - \tilde{g}|) \\
&\quad + \mu(|\tilde{g}|)\mu(|f - \tilde{f}|) + |\mu(\tilde{f} \circ T^n \tilde{g}) - \mu(\tilde{f})\mu(\tilde{g})| \\
&\leq 2\epsilon(\|\tilde{f}\|_{L^2(\mathfrak{m})} + \|\tilde{g}\|_{L^2(\mathfrak{m})}) + |\mu(\tilde{f} \circ T^n \tilde{g}) - \mu(\tilde{f})\mu(\tilde{g})| \\
&< 2\epsilon(\|\tilde{f}\|_{L^2(\mathfrak{m})} + \|\tilde{g}\|_{L^2(\mathfrak{m})} + 1),
\end{aligned}$$

when n is large enough. Remark, that we used Cauchy-Schwartz inequality extensively to bound L^1-norms.

Finally, we wish to point out that the mixing measure we found is the only ergodic measure that has a density (w.r.t. Lebesgue measure). First, there is a general result

that ergodic measures are mutually singular [6, p. 152]. Additionally, Corollary A.1 in the Appendix states that the domain of the mixing measure covers the whole circle. By noticing that mixing measures are ergodic, together these theorems imply that if there is another ergodic measure, then its domain has a zero Lebesgue measure. That is, the measure is not absolutely continuous w.r.t. Lebesgue measure.

Appendix A.

Two theorems, which implement very central properties of expanding dynamics, are presented here since fitting them smoothly inside the treatise turned out to be difficult. Nevertheless, that does not diminish their value.

Theorem A.1. *Suppose $\psi \in C^\alpha$ (for fixed α). Then $\mathcal{L}^n\psi \in C^\alpha$ for every $n \in \mathbb{N}$ and its Hölder constant satisfies*

$$H_\alpha(\mathcal{L}^n\psi) \leq \left(\frac{H_\alpha(\psi)}{\lambda^\alpha} + (\exp(\Omega)-1)(H_\alpha(\psi)+1)\right)(1+\Omega). \tag{32}$$

Proof. Let $|x-y|$ be small enough so that we can use the estimate from lemma 4.1. Then it holds that

$$\begin{aligned}
&|\mathcal{L}^n\psi(x) - \mathcal{L}^n\psi(y)| \\
&= \left|\sum_{i=1}^k \frac{\psi((T^n)_i^{-1}(x))}{(T^n)'((T^n)_i^{-1}(x))} - \frac{\psi((T^n)_i^{-1}(y))}{(T^n)'((T^n)_i^{-1}(y))}\right| \\
&\leq \sum_{i=1}^k \frac{|\psi((T^n)_i^{-1}(x)) - \psi((T^n)_i^{-1}(y))|}{(T^n)'((T^n)_i^{-1}(x))} + \sum_{i=1}^k \left|\frac{\psi((T^n)_i^{-1}(y))}{(T^n)'((T^n)_i^{-1}(x))} - \frac{\psi((T^n)_i^{-1}(y))}{(T^n)'((T^n)_i^{-1}(y))}\right| \\
&\leq H_\alpha(\psi)\sum_{i=1}^k \frac{d((T^n)_i^{-1}(x),(T^n)_i^{-1}(y))^\alpha}{(T^n)'((T^n)_i^{-1}(x))} + \sum_{i=1}^k \frac{|\psi((T^n)_i^{-1}(y)|}{(T^n)'((T^n)_i^{-1}(y))}\left|\frac{(T^n)'((T^n)_i^{-1}(y))}{(T^n)'((T^n)_i^{-1}(x))} - 1\right| \\
&\leq \frac{H_\alpha(\psi)}{\lambda^{\alpha n}}\sum_{i=1}^k \frac{d(x,y)^\alpha}{(T^n)'((T^n)_i^{-1}(x))} + \|\psi\|_\infty \sum_{i=1}^k \frac{1}{(T^n)'((T^n)_i^{-1}(y))}|e^{\Omega|x-y|}-1| \\
&\leq \frac{H_\alpha(\psi)}{\lambda^{\alpha n}}\|\mathcal{L}^n\mathbf{1}\|_\infty d(x,y)^\alpha + \|\psi\|_\infty\|\mathcal{L}^n\mathbf{1}\|_\infty|e^{\Omega d(x,y)}-1| \\
&\leq \left(\frac{H_\alpha(\psi)}{\lambda^\alpha} + (\exp(\Omega)-1)\|\psi\|_\infty\right)\|\mathcal{L}^n\mathbf{1}\|_\infty d(x,y)^\alpha
\end{aligned}$$

Additionally, for probability density ψ there exists x_0 such that $\psi(x_0) = 1$. Thus, for every x it holds that

$$|\psi(x)| = |\psi(x) - \psi(x_0) + \psi(x_0)| \leq H_\alpha(\psi)d(x,x_0)^\alpha + 1 \leq H_\alpha(\psi) + 1.$$

Combining this with the uniform bound for $\|\mathcal{L}^n\mathbf{1}\|_\infty$ results

$$H_\alpha(\mathcal{L}^n\psi) \leq \left(\frac{H_\alpha(\psi)}{\lambda^\alpha} + (\exp(\Omega)-1)(H_\alpha(\psi)+1)\right)(1+\Omega). \tag{33}$$

□

Theorem A.2. *Suppose ψ is a Hölder continuous probability density. Then there exist $a > 0$ and $N_1 = N_1(\psi) \in \mathbb{N}$ with the property that $\inf_x \mathcal{L}^n\psi(x) \geq a$ for all $n \geq N_1$.*

Proof. First, remark that, by previous theorem, there is $L \in \mathbb{R}_+$ such that $H_\alpha(\mathcal{L}^n\psi) \leq L$ for every $n \in \mathbb{N}$.

Let's start with ψ. Since its probability density, there is a point $x_0 \in \mathbb{S}^1$ such that $\psi(x_0) = 1$. Using the fact that $H_\alpha(\psi) \leq L$ we have

$$d(x, x_0) < \left(\frac{1}{2L}\right)^{1/\alpha} \Longrightarrow |\psi(x) - 1| = |\psi(x) - \psi(x_0)| \leq Ld(x, x_0)^\alpha < \frac{1}{2} \Longrightarrow \psi(x) \geq \frac{1}{2}$$

More briefly:

$$x \in B(x_0, (1/(2L))^{1/\alpha}) \Longrightarrow \psi(x) \geq \frac{1}{2}.$$

If $L \leq 2^{1-\alpha}$ the claim holds for every $a \in [0, 1/2]$. Thus, we may assume $L > 2^{1-\alpha}$.

Because of the expanding property of T, there is some N_1 such that

$$1 < \lambda^{N_1} |B(x_0, (1/(2L))^{1/\alpha})|$$

and furthermore $T^{N_1}(B(x_0, (1/(2L))^{1/\alpha})) = \mathbb{S}^1$. This is the case when

$$\lambda^{N_1} > (2L)^{1/\alpha} > 1/|B(x_0, 1/(2L)^{1/\alpha})| \text{ i.e. } N_1 > \frac{\log(2L)}{\alpha \log(\lambda)}.$$

Now set $N_1 = 1 + \lceil \log(2L)/(\alpha \log(\lambda)) \rceil$. (We write $\lceil x \rceil$ for the smallest integer $\geq x$.) By the fact that the preimage $T^{-N_1}x$ of every point $x \in \mathbb{S}^1$ intersects $B(x_0, 1/(2L)^\alpha)$, we have that

$$\mathcal{L}^{N_1}\psi(x) = \sum_{y \in T^{-N_1}\{x\}} \frac{\psi(y)}{(T^{N_1})'(y)} \geq \frac{1}{2\|(T^{N_1})'\|_\infty} \geq \frac{1}{2\|T'\|_\infty^{N_1}}.$$

Additionally, since $H_\alpha(\mathcal{L}^{n-N_1}\psi) \leq L$ for *every* $n > N_1$, we can do the same reasoning and conlclude that

$$\mathcal{L}^n\psi(x) = \mathcal{L}^{N_1}(\mathcal{L}^{n-N_1}\psi)(x) \geq \frac{1}{2\|T'\|_\infty^{N_1}}. \tag{34}$$

This finishes the proof (by definition of infimum). □

The following corollary is immediate.

Corollary A.1. *The invariant density ϕ found earlier is strictly positive.*

References

[1] William Ott, Mikko Stenlund, and Lai-Sang Young, *Memory loss for time-dependent dynamical systems*, Mathematical Research Letters, 16(3):463-475, 2009.

[2] Carlangelo Liverani, Lecture notes for the course "Sistemi Dinamici" (2011-2012) [pdf]. Available at: <www.mat.uniroma2.it/~liverani/SysDyn10/chap6.pdf> [Accessed 11 November 2012].

[3] Lai-Sang Young, *Ergodic Theory of Differentiable Dynamical Systems* [pdf], 1995. Available at: <www.math.nyu.edu/~lsy/papers/lecturenotes93.pdf> [Accessed 11 November 2012]. Also published in "Real and Complex Dynamical Systems", ed. Branner and Hjorth, NATO ASI series, Kluwer Academic Publishers, (1995).

[4] Torgny Lindvall, *Lectures on the Coupling Method*, Wiley, New Yourk, 1992.

[5] Olle Häggström, *Finite Markov Chains and Algorithmic Applications*, Cambridge University Press, 1st edition, 2002.

[6] Peter Walters, *An Introduction to Ergodic theory*, Springer, 1st softcover edition, 2000.

(Henri Sulku) Department of Mathematics and Statistics, P.O. Box 68, Fin-00014 University of Helsinki, Finland.

E-mail address: henri.sulku@helsinki.fi